What is claimed is:

--1. A method for monitoring middleware performance, the method comprising:

determining a route time for a message transmitted

5 along a predetermined network route;

determining at least one queue residency time, the at least one queue residency time reflecting an amount of time at least one other message is stored in at least one respective queue located along the predetermined network route; and

calculating a middleware response time according to the route time and the at least one queue residency time.

15 --2. The method as set forth in claim 1, wherein determining the route time includes determining an amount of time for the message to travel from a source node along the predetermined network route to a destination node and then back to the source node.

20

10

--3. The method as set forth in claim 2, wherein determining the route time includes determining an origination timestamp indicating when the message was sent from the source node,

storing the origination timestamp in a field of the message,

determining an end timestamp indicating when the source node receives the message from the destination node,

5

storing the end timestamp in another field of the message, and

calculating the route time by determining the difference between the origination timestamp and the end timestamp.

- --4. The method as set forth in claim 1, wherein determining the at least one queue residency time includes
- sampling a plurality of application messages,

 determining an identification for the sampled

 message, the identification being stored in a field of
 the sampled message,

storing the identification of the sampled message,

determining a put timestamp for the sampled

message, the put timestamp indicating when the sampled

message was placed on a local production queue,

comparing identifications associated with messages retrieved from the local production queue with the identification of the sampled message,

determining a get timestamp for the sampled message, if a match is determined, the get timestamp indicating when the sampled message was retrieved from the local production queue, and

calculating a queue residency time by determining the difference between the put timestamp and the get timestamp.

- --5. The method as set forth in claim 1, wherein calculating the middleware response time includes adding the route time and the at least one queue residency time.
 - --6. The method as set forth in claim 1, wherein the message is a user-defined sample message replicating an actual application message.

20

5

--7. The method as set forth in claim 6, wherein the user-defined sample message is a plurality of messages batched together, the plurality of messages having the same or varying sizes.

--8. A method for monitoring middleware performance, the method comprising:

determining a route time for a user-defined sample message to be transmitted along a predetermined network route, the route time reflecting an amount of time for the sample message to travel from an origination queue manager to a destination queue manager and then back along the same route to the origination queue manager;

determining at least one queue residency time for at least one local production queue, the at least one local production queue being associated with the origination queue manager and/or the destination queue manager and the at least one queue residency time reflecting the amount of time an actual application message is stored in the at least one local production queue; and

calculating a middleware response time by adding the route time to the at least one queue residency time.

20

5

10

15

- --9. A system for monitoring middleware performance, the system comprising:
- a computer system adapted to generate a sample message and an application message; and

a computer memory electrically connected to the computer system encoded with instructions for performing the following:

determining a route time for the sample message

5 transmitted along a predetermined network route;

10

determining at least one queue residency time, the at least one queue residency time reflecting an amount of time the application message is stored in at least one respective queue located along the predetermined network route; and

calculating a middleware response time according to the route time and the at least one queue residency time.